

**Autor**

Manuel Blázquez Ochando  
Facultad de Ciencias de la Documentación  
Departamento de Biblioteconomía y Documentación  
[manuel.blazquez@pdi.ucm.es](mailto:manuel.blazquez@pdi.ucm.es)

**Título**

Análisis webométrico de los medios de comunicación brasileños: prensa, radio y televisión

**Resumen**

El objeto de la investigación es el análisis cuantitativo de la web brasileña de medios de comunicación, prensa, radio y televisión, obtenidos mediante técnicas de webcrawling, que en suma representan 1315 entidades informativas. La metodología empleada, propia de la webmetría, ya aplicada en otros estudios (GRAELLS, E. and Baeza Yates, R., 2007), permite conocer el grado de desarrollo de la web, su nivel de interconexión, su macroestructura, formatos, extensiones, dominios, páginas más enlazadas, cantidad de documentos, tamaño y recursos de utilidad para su posterior aprovechamiento y reutilización. Finalmente se propone una segunda fase del estudio en la que se llevan a cabo investigaciones cualitativas sobre la producción informativa y su clasificación temática, partiendo de los recursos recopilados en ésta primera aproximación.

## 1. Planteamiento y objetivos

El objetivo de la investigación es el estudio de los medios de comunicación brasileños con un enfoque cuantitativo de su webmetría y cualitativo sobre su producción informativa. Dada la amplitud y complejidad del trabajo, el presente documento, muestra la primera parte de la investigación en la que se describen las características de la web de prensa, radio y televisión de Brasil.

Existe por otra parte una doble motivación para el desarrollo del presente estudio, que justifica su realización. Por un lado la celebración del “*I Seminario hispano-brasileño de investigación en información, documentación y sociedad*”, que brinda la posibilidad de abrir una nueva línea de investigación y cooperación entre especialistas. En segundo lugar, la reciente retirada de los principales diarios brasileños del agregador de noticias Google News, noticia que fue cubierta por múltiples medios de comunicación en todo el mundo, dada su repercusión, como por ejemplo la cadena británica (BBC, 2012), el diario digital de economía *PaidContent* (ANDREWS, R., 2012), el portal de noticias económicas de *Mashable* (INDVIK, L., 2012), el sitio web de noticias globales en línea *Globalpost* (RALPH, T., 2012) o el diario digital de tecnología *The Verge* (WEBSTER, A., 2012). Según se informa, los medios de prensa brasileños decidieron prohibir la agregación de información de sus diarios digitales a Google debido a la pérdida de tráfico y circulación de usuarios en sus sitios web, lo que implicaba un estancamiento en el crecimiento del número de lectores y por tanto de sus propios clientes. Esta situación plantea diversas cuestiones en relación a la capacidad de los sistemas de agregación de noticias y portales de contenidos multicanales que toman como fuente de datos los canales de información públicos de terceros medios de comunicación. Por otra parte, justifica la necesidad de conocer cuál es el estado de la cuestión en los medios de comunicación digitales de Brasil, qué características tienen, cuántos canales de sindicación son susceptibles de agregación, qué información se publica en ellos, hasta qué punto puede ser aprovechada, cuál es su tamaño y relevancia, qué visibilidad tienen a través del uso de metadatos, cuáles son los medios más enlazados y más relevantes por extensión, cantidad de documentos, entre otros interrogantes.

## 2. Desarrollo de la investigación

En la primera fase de la investigación, se emplea una metodología basada en el análisis webométrico de los medios de comunicación digitales de Brasil, empleando como guía estudios de naturaleza similar como por ejemplo (BAEZA YATES, R. et al., 2006) y (GRAELLS, E.; Baeza Yates, R., 2007). Para ello es necesario elaborar un listado exhaustivo de las direcciones URL de los medios de comunicación de prensa, radio y televisión con presencia digital, partiendo de las fuentes de información disponibles en la web. Los medios de prensa fueron obtenidos a través del *Instituto Verificador de Circulação* (IVC, 2012), organismo regulador del sector; los medios de radio mediante el directorio especializado *Radio Station World*, (HONKENSEN, T.C., 2012); y la televisión, con el directorio de medios (MEDIA ONLINE, 2012). De su análisis se extraen un total de 340 medios de prensa, 740 medios de radio y 241 medios de televisión, sumando un total de 1.315 medios de comunicación. La extracción de datos y la tabulación de los resultados es realizada automáticamente con la herramienta webcrawler Mbot, divulgada en anteriores trabajos, véase (BLÁZQUEZ OCHANDO, M., 2011) y (BLÁZQUEZ OCHANDO, M.; Serrano Mascaraque, E., 2011), capaz de generar una relación de datos completa sobre el número total de enlaces de cada medio, distinción entre sitios y páginas, distribución de dominios enlazados, tipología de documentos, macroestructura de la web, extracción de metadatos, canales de sindicación y rankings.

En la segunda fase de la investigación, que será presentada en sucesivas publicaciones, se estudiará la producción informativa de los medios de comunicación brasileños a partir de los contenidos publicados en sus canales de sindicación, previamente recopilados. El análisis cualitativo de la información comprende su clasificación temática automática a través de la plataforma para la investigación de canales de sindicación *Resync*, ya utilizada en anteriores investigaciones sobre la producción informativa de los medios de comunicación españoles y mexicanos (BLÁZQUEZ OCHANDO, M., 2012). En tal caso el método comprende la selección del vocabulario de clasificación, su adaptación, preparación de los contenidos informativos recopilados, su normalización, indexación y ejecución de los algoritmos de clasificación del programa.

### 3. Resultados

Utilizando los listados de los medios de comunicación como semilla inicial del estudio, el programa Mbot se configura para efectuar un proceso de extracción y análisis de datos con una profundidad de dos niveles de enlazamiento. La profundidad del análisis es suficiente para obtener una imagen muy aproximada de las características de la web de medios de comunicación brasileños, puesto que en total han sido analizados cerca de medio millón de enlaces únicos, véase *tabla1*. Esto significa que todos fueron obtenidos desestimando las duplicaciones correspondientes, ya que en tal caso la cifra ascendería a un total de 2.022.892 enlaces, dificultando o añadiendo ruido a la precisión de los resultados. Según se observa y comprueba, durante los distintos niveles de análisis, el sector de medios de prensa es el más voluminoso tanto en cantidad de sitios web como por páginas. Este hecho resulta paradójico si se tiene en cuenta que el número de medios de radio analizados dobla a los de prensa. Ello se explica debido a una mayor cantidad de documentación textual, así como de páginas que contienen las noticias de prensa, en contraposición al modelo de radio web que prima los servicios de radiodifusión digital, viéndose reducido el número de páginas destinadas a soportar noticias textuales. No obstante, existen importantes excepciones, debido a la presencia de grandes conglomerados multisectoriales que abordan indistintamente en sus webs prensa, radio y televisión. Estos grupos están muy presentes en el estudio y son Opopo, UOL, Clic RBS, EBC y Globo.

Datos generales del análisis								
Categoría	Nº de medios	Nº total de enlaces únicos analizados	Nivel 1			Nivel 2		
			Enlaces únicos analizados	Sitios	Páginas	Enlaces únicos analizados	Sitios	Páginas
Prensa	340	204.133	18.468	2.243	16.225	185.665	15.114	170.551
Radio	734	133.322	12.771	2.235	10.536	120.551	11.566	108.985
Televisión	241	146.257	9.791	1.007	8.784	136.466	6.231	130.235
Total	1.315	483.712	41.030	5.485	35.545	442.682	32.911	409.771

*Tabla 1. Datos generales del análisis webométrico sobre medios de comunicación brasileños*

En cuanto a la distribución de dominios de nivel superior, también denominados TLD, destacan los dominios genéricos de tipo comercial *.com* con una presencia media cercana al 90% en todos los medios por encima de los de tipo *.org* y *.net*. Apenas tienen presencia los dominios académicos *.edu* y gubernamentales *.gov*, lo que indica que apenas son enlazados por los medios de comunicación, véase *tabla2*. Por otra parte los dominios de tipo geográfico, muestran un evidente predominio del dominio brasileño *.br* con más de 142.060 páginas enlazadas desde los sitios web de prensa, 85.832 desde los sitios web de radio y 56.371 desde la televisión. También resulta significativo que los dominios geográficos extranjeros más vinculados sucesivamente sean los españoles *.es*, bolivianos *.bo* y estadounidenses *.us*.

Distribución de dominios de nivel superior TLD							
TLD Genéricos							
TLD		Prensa		Radio		Televisión	
		Sitios	Páginas	Sitios	Sitios	Páginas	Sitios
.com	Comercial	2.069	44.052	1.758	29.812	1.452	75.858
.org	Organizaciones	114	147	93	293	96	1.575
.net	Red	273	3.013	113	146	74	128
.edu	Académico	5	4	2	1	4	3
.gov	Gubernamental	15	10	2	1	0	2
Subtotal		2.476	47.226	1.968	30.253	1.626	77.566
TLD Geográficos							
TLD		Prensa		Radio		Televisión	
		Sitios	Páginas	Sitios	Sitios	Páginas	Sitios
.de	Alemania	9	12	6	3	7	3
.ar	Argentina	0	8	0	6	0	12
.bo	Bolivia	0	239	0	135	0	311
.br	Brasil	4.355	142.060	4.704	85.832	2.246	56.371
.cl	Chile	0	0	0	5	0	4
.co	Colombia	0	0	0	0	0	0
.ec	Ecuador	0	0	0	1	0	0
.es	España	2	6	36	3.046	22	1.211
.us	Estados Unidos	12	884	4	23	92	98
.fr	Francia	6	8	5	1	10	3
.it	Italia	3	2	5	4	15	2
.mx	México	29	1	25	3	38	1
.py	Paraguay	0	1	0	2	0	9
.pe	Perú	0	2	0	2	0	3

.pt	Portugal	0	8	0	2	0	79
.uk	Reino Unido	14	47	11	18	10	13
.uy	Uruguay	0	0	0	0	0	2
.ve	Venezuela	0	0	0	0	0	0
Subtotal		4.430	143.278	4.796	89.083	2.440	58.122

Tabla 2. Distribución de dominios TLD en los medios de comunicación brasileños

Sobre la tipología de los documentos en la web analizada, se destaca que la web de medios de comunicación brasileños utilizan un mayor número de páginas web estáticas que dinámicas, véase *tabla3*. De hecho si se suma la cantidad de páginas dinámicas en formato *.php* y *.asp*, no se alcanza en ningún caso la mitad del total de páginas estáticas en *.html*. Ello significa que aún se está lejos de una modernización global de toda la web de medios de comunicación brasileños, que no todos los sectores emplean gestores de contenidos. A pesar de ello, la cantidad de webs dinámicas, indica que existe un sector que sí trabaja de acuerdo a estas tecnologías, que coincide con los grandes grupos de comunicación señalados anteriormente.

Tipología de documentos en la web											
Formatos de páginas web											
	html		php		asp		xml				
Prensa	45.143		13.075		5.433		330				
Radio	25.132		10.067		2.923		280				
Televisión	31.667		6.338		3.375		190				
Formatos de documentos											
	pdf		doc		ppt		xls		rtf		mdb
Prensa	380		24		2		2		0		0
Radio	660		91		10		5		0		0
Televisión	544		85		8		5		0		0
Formatos de audio y vídeo											
	swf	avi	mpg	mp4	mov	wmv	mp3	wav	wma	ogg	
Prensa	9	0	0	1	0	0	48	5	0	0	
Radio	4	0	0	1	0	0	470	0	1	0	
Televisión	4	0	0	13	1	379	180	0	0	0	
Formatos de imagen											
	png		jpg		gif		tif		svg		bmp
Prensa	5.920		42.215		4.031		1		0		46
Radio	5.223		29.017		2.649		17		0		35
Televisión	3.429		30.619		1.975		0		0		27

Tabla 3. Tipos de documentos según su extensión y formato en la web de medios de comunicación brasileños

En cuanto a los formatos de documento más populares, se destaca el *.pdf*, en menor medida el *.doc* y de forma casi inapreciable las hojas de cálculo y presentaciones. Las cifras obtenidas en todo caso son bajas, lo que indica que los medios de comunicación brasileños no suelen publicar gran cantidad de documentos ofimáticos en sus webs, empleando en tal caso páginas estáticas de tipo *.html*. Ello permitiría también explicar el tamaño de la web estática por encima de la dinámica. En relación a los formatos de audio y video se observa que la prensa no suele cargar en sus servidores este tipo de archivos. En cambio la radio y televisión utilizan más archivos de audio *.mp3* y video *.wmv* respectivamente. Finalmente los formatos de imagen más utilizados por los medios de prensa, radio y televisión, son los *.jpg* con cifras superiores a las 30.000 imágenes, *.png* que varían entre las 3.000 y 6.000 y *.gif* empleadas habitualmente en el interfaz gráfico de los sitios web, entre las 2.000 y 4.000.

Seguidamente el webcrawler Mbot obtuvo los siguientes datos de la macroestructura de los distintos sectores de prensa, radio y televisión, véase *tabla4*. Se observa que el núcleo de enlaces fuertemente conectados en el componente *Main* varía entre el 3,3% y el 5,2% del total. En suma, junto con los componentes *In*, *Out*, *Tentacle In*, *Tentacle Out* y *Tunnel*, representan un 9,36% de los enlaces de prensa, un 10,95% de los enlaces de radio y un 7,77% de los enlaces de televisión. Esta información permite advertir que más del 80% de la web está pobremente interconectada entre sí, planteando problemas de visibilidad y un ranking reducido en motores de búsqueda. De hecho el componente débilmente vinculado de la web varía entre el 85% de la radio y el 89% de la televisión, quedando reflejado en el componente *Island*. También se debe reseñar, que el programa Mbot no pudo analizar cerca de un 3% de enlaces de la macroestructura lo que significa que existe una desviación por error y que los datos pueden no ser totalmente exactos. No obstante, se ha constatado que entre el 96% y el 97% de los enlaces de prensa, radio y televisión fueron completamente analizados.



Tabla 4. Análisis de la macroestructura de la web de medios de comunicación brasileños

Los medios de comunicación con más metadatos son los que se muestran en la *tabla5*, destacando en el rango de los 10 primeros, algunos de los grupos de comunicación más importantes de Brasil, concretamente en prensa destaca *UOL* con más de 2.800 metadatos, seguido del diario digital *Araraquara.com* con 1.700 y el portal de prensa y noticias *Clic RBS* con cifras similares. Resulta interesante la presencia del grupo Opovo

y Globo, tal como se demostrará en los sucesivos rangos de datos. La suma de los metadatos de los diez primeros medios de la lista, suponen el 41% del total, lo que indica que los 330 medios de prensa restantes acaparan el 69% de los metadatos. En cuanto a la radio, las cifras son más reducidas si se comparan con la prensa, destacando *GloboRadio.com* con 648, *Jovem Pan Online* del grupo UOL con 395 y *Radio educativa fm* de la Prefeitura Municipal de Campinas con 355. Los diez primeros medios de la radio con más metadatos, suman un total de 2.923 lo que supone el 27% del total registrado en el sector. Finalmente, los metadatos en televisión tienen unos valores ligeramente superiores a la radio. Destacan el canal *Band TV* del grupo UOL con 766 metadatos, *SporTV* del grupo Globo con 412 y *Brasil TV* de la EBC con 371. Los diez primeros medios de televisión brasileños indicados en el listado, aglutinan el 36% de todos los metadatos del sector, lo que igualmente supone un factor relevante de cara a su visibilidad y posicionamiento en la web.

Medios de comunicación con más metadatos			
Prensa			
Rango	Medio	URL	Nº de metadatos
1	UOL o Melhor Conteúdo	<a href="http://www.uol.com.br">http://www.uol.com.br</a>	2.815
2	Araraquara.com	<a href="http://www.araraquara.com">http://www.araraquara.com</a>	1.772
3	Clic RBS	<a href="http://www.clicrbs.com.br">http://www.clicrbs.com.br</a>	1.712
4	Opovo online	<a href="http://www.opovo.com.br">http://www.opovo.com.br</a>	1.000
5	Otempo online	<a href="http://www.otempo.com.br">http://www.otempo.com.br</a>	545
6	Jornal do Oeste	<a href="http://www.jornaldooeste.com.br">http://www.jornaldooeste.com.br</a>	434
7	Folha de S.Paulo	<a href="http://www.folha.uol.com.br">http://www.folha.uol.com.br</a>	406
8	G1 o Portal de noticias da Globo	<a href="http://g1.globo.com">http://g1.globo.com</a>	398
9	Bol	<a href="http://www.bol.uol.com.br">http://www.bol.uol.com.br</a>	395
10	Jornal o Globo	<a href="http://oglobo.globo.com">http://oglobo.globo.com</a>	391
Subtotal			9.868
Total			24.363
Radio			
Rango	Medio	URL	Nº de metadatos
1	GloboRadio.com	<a href="http://globoradio.globo.com">http://globoradio.globo.com</a>	648
2	Jovem pan Online	<a href="http://jovempan.uol.com.br">http://jovempan.uol.com.br</a>	395
3	Educative fm	<a href="http://www.educativa.fm.br">http://www.educativa.fm.br</a>	355
4	EBC Rádio	<a href="http://www.ebc.com.br">http://www.ebc.com.br</a>	349
5	Itatiaia	<a href="http://www.itatiaia.com.br">http://www.itatiaia.com.br</a>	310
6	Rádio MEC	<a href="http://radiomec.com.br">http://radiomec.com.br</a>	237
7	Rádio Cruz de Malta	<a href="http://www.radiocruzdemalta.com.br">http://www.radiocruzdemalta.com.br</a>	212
8	96fm marapiraca	<a href="http://www.96fmarapiraca.com.br">http://www.96fmarapiraca.com.br</a>	151
9	Radio UNESP	<a href="http://radio.unesp.br">http://radio.unesp.br</a>	133
10	Radio Cascavel	<a href="http://www.cbncascavel.com.br">http://www.cbncascavel.com.br</a>	133
Subtotal			2.923
Total			10.895
Televisión			
Rango	Medio	URL	Nº de metadatos
1	Band news TV	<a href="http://www.band.com.br">http://www.band.com.br</a>	766
2	SporTV	<a href="http://sportv.globo.com">http://sportv.globo.com</a>	412
3	EBC TV Brasil	<a href="http://tvbrasil.ebc.com.br/">http://tvbrasil.ebc.com.br/</a>	371
4	Fox Sports	<a href="http://www.foxsports.com.br">http://www.foxsports.com.br</a>	341
5	Jangadeiro TV	<a href="http://www.jangadeiroonline.com.br">http://www.jangadeiroonline.com.br</a>	320
6	Rede Globo TV	<a href="http://redegloboglobo.com">http://redegloboglobo.com</a>	305
7	Rede Super de Televisão	<a href="http://redesuper.com.br">http://redesuper.com.br</a>	287
8	Just TV	<a href="http://www.justtv.com.br">http://www.justtv.com.br</a>	248
9	Mix TV	<a href="http://www.mixtv.com.br">http://www.mixtv.com.br</a>	214
10	Rede Correio TV	<a href="http://www.rc.tv.br">http://www.rc.tv.br</a>	197
Subtotal			3.461
Total			9.631

Tabla 5. Medios de comunicación brasileños con más enlaces únicos (enlaces no duplicados)

Los medios de comunicación con más enlaces únicos, en muchos casos coinciden con los que más metadatos tienen, ello puede comprobarse en la *tabla6*. Se observa que los grupos de comunicación de prensa *UOL*, *Clic RBS* y *Opovo* ocupan los primeros puestos de la lista con más de 10.000 enlaces únicos contabilizados. El subtotal de los diez primeros medios de prensa con más enlaces supone el 64% del total de enlaces únicos analizados en el sector, lo que indica una clara descompensación de tamaño entre los principales grupos periodísticos y los más reducidos. En radio destacan *Jovem Pan*, *GloboRadio.com* y *EBC Rádios* con un rango variable entre los 6.443 y los 2.367 enlaces únicos respectivamente. En televisión los más destacados son *SporTV* con 5.920 enlaces, seguido de *Band TV* con 3.899 y *Jangadeiro TV* con 3.858. También se observa que no siempre los medios de comunicación con más enlaces únicos tienen más dominios, por ejemplo la radio *Jovem Pan* que consta de más enlaces únicos, dispone de 124 dominios, en contraposición de los 2.741 que tiene *GloboRadio.com*, la segunda radio reseñada en la lista.

Medios de comunicación con más enlaces únicos					
Prensa					
Rango	Medio	URL	Nº total enlaces únicos	Nº dominios enlazados	Nº páginas enlazadas
1	UOL o Melhor Conteúdo	http://www.uol.com.br	31.826	5.056	26.770
2	Clic RBS	http://www.clicrbs.com.br	25.309	568	24.741
3	Opovo online	http://www.opovo.com.br	13.760	215	13.545
4	Araraquara.com	http://www.araraquara.com	11.000	568	10.432
5	Folha de S.Paulo	http://www.folha.uol.com.br	6.810	287	6.523
6	G1 Noticias da Globo	http://g1.globo.com	6.787	156	6.631
7	Estadão.com	http://www.estadao.com.br	5.055	1.337	3.718
8	Bol	http://www.bol.uol.com.br	4.252	321	3.931
9	Lancenet.com	http://www.lancenet.com.br	3.709	46	3.663
10	Jornal o Globo	http://oglobo.globo.com	3.592	530	3.062
Subtotal			112.100	9.084	103.016
Total			173.820		
Radio					
Rango	Medio	URL	Nº total enlaces únicos	Nº dominios enlazados	Nº páginas enlazadas
1	Jovem pan Online	http://jovempan.uol.com.br	6.443	124	6.319
2	GloboRadio.com	http://globoradio.globo.com	4.975	2.741	2.234
3	EBC Radio	http://www.ebc.com.br	2.367	22	2.345
4	Educativa fm	http://www.educativa.fm.br	1.636	996	640
5	Studio fm 99.1	http://www.studiofm.com.br	1.218	19	1.199
6	Stylo 102.1 fm	http://www.stylofm.com.br	1.216	21	1.195
7	Radio 7	http://www.r7.com	1.028	130	898
8	94 fm	http://www.94fm.com.br	919	7	912
9	Mundial fm	http://www.fmmundi.com.br	874	10	864
10	93 fm	http://radio93.com.br/	740	12	728
Subtotal			21.416	4.082	17.334
Total			55.641		
Televisión					
Rango	Medio	URL	Nº total enlaces únicos	Nº dominios enlazados	Nº páginas enlazadas
1	SporTV	http://sportv.globo.com	5.920	330	5.590
2	Band news TV	http://www.band.com.br	3.899	326	3.573
3	Jangadeiro TV	http://www.jangadeiroonline.com.br/tv/	3.858	61	3.797
4	TV Novo Tempo	http://novotempo.com	2.855	85	2.770
5	EBC Brasil TV	http://www.ebc.com.br	2.717	54	2.663
6	Rede Globo TV	http://redeglobo.globo.com	1.820	117	1.703
7	Rede Super de Televisão	http://www.redesuper.com.br	1.801	45	1.756
8	Mix TV	http://www.mixtv.com.br	1.267	7	1.260
9	Agência de Notícias do Acre	http://www.videos.ac.gov.br	1.246	14	1.232
10	IRDEB TV	http://www.irdeb.ba.gov.br	1.003	28	975
Subtotal			26.386	1.067	25.319
Total			48.143		

Tabla 6. Medios de comunicación brasileños con más enlaces únicos (enlaces no duplicados)

Los medios de comunicación con más documentos y archivos, incluyendo imágenes, audio y vídeo, siguen siendo los principales grupos de comunicación anteriormente referidos, véase *tabla7*. De hecho el 64% de los documentos y archivos disponibles en los medios de prensa analizados corresponden a los diez primeros de la lista. En el caso de los medios televisivos analizados, el porcentaje disminuye al 51% y en el sector de radio al 33%.

Medios de comunicación con más documentos y archivos			
Prensa			
Rango	Medio	URL	Nº total de documentos únicos
1	UOL o Melhor Conteúdo	http://www.uol.com.br	12.421
2	Clic RBS	http://www.clicrbs.com.br	6.900
3	Opovo online	http://www.opovo.com.br	3.584
4	G1 o Portal de noticias da Globo	http://g1.globo.com	3.309
5	Bol	http://www.bol.uol.com.br	2.781
6	Estadão.com	http://www.estadao.com.br	1.765
7	Jornal o Globo	http://oglobo.globo.com	1.258
8	Folha de S.Paulo	http://www.folha.uol.com.br	1.243
9	Araraquara.com	http://www.araraquara.com	1.161
10	Lancenet.com	http://www.lancenet.com.br	1.034
Subtotal			35.456
Total			55.523



Radio			
Rango	Medio	URL	Nº total de documentos únicos
1	Jovem pan Online	http://jovempan.uol.com.br	1.875
2	GloboRadio.com	http://globoradio.globo.com	1.690
3	EBC Radio	http://www.ebc.com.br	557
4	94 fm	http://www.94fm.com.br	528
5	Educativa fm	http://www.educativa.fm.br	431
6	98 fm Curitiba	http://www.98fmcuitiba.com.br	407
7	Radio 7	http://www.r7.com	315
8	Culturabrasil portal	http://www.culturabrasil.com.br	278
9	Alpha fm	http://www.alphafm.com.br	276
10	Mundial fm	http://www.fmmundi.com.br/	184
Subtotal			6.541
Total			19.628
Televisión			
Rango	Medio	URL	Nº total de documentos únicos
1	SporTV	http://sportv.globo.com	2.346
2	Band news TV	http://band.com.br	1.609
3	Jangadeiro TV	http://www.jangadeiroonline.com.br	1.385
4	Rede Super de Televisão	http://www.redesuper.com.br	968
5	TV Novo Tempo	http://novotempo.com/tv	834
6	Rede Globo TV	http://redegloboglobo.com	795
7	EBC TV Brasil	http://www.ebc.com.br	648
8	Boa Vontade TV	http://www.boavontade.com	521
9	Mix TV	http://www.mixtv.com.br	495
10	Fox Sports	http://www.foxsports.com.br	454
Subtotal			10.055
Total			19.588

Tabla 7. Medios de comunicación brasileños con más documentos y archivos (incluyendo imágenes, audio y vídeo)

Seguidamente, el webcrawler Mbot realiza el proceso de análisis y extracción de los canales de sindicación en todos los medios de comunicación brasileños, véase *tabla8*. Estos canales de sindicación serán utilizados para realizar investigaciones posteriores, especialmente el análisis de la producción informativa de los medios a través de las noticias publicadas en dichos canales. El diario digital con más canales de sindicación únicos es *Araraquara.com* con 545, seguido de los grupos *UOL* con 264 y *Opovo* con 230.

En total, la prensa brasileña consta de 2.730 canales de sindicación, pero los diez primeros medios de la lista aglutinan el 69% de tales canales, confirmando las tendencias anteriormente relatadas y la desigualdad o falta de equilibrio del sector. En cuanto a radio y televisión los primeros diez medios con más canales de sindicación conforman el 63% y el 70% de todos los canales de sindicación para su sector respectivamente. En suma la información producida por los medios de comunicación digitales brasileños podría obtenerse a través de 4.509 canales de sindicación. Google News Brasil utiliza sólo 1.500 canales de los que se abastece (GOOGLE, 2012a), para alimentar un portal de contenidos con el que compite y logra reducir el tráfico de los principales medios de comunicación brasileños. Esto supone, que se ha logrado extraer más de 3.000 canales de sindicación que Google no está empleando.

Medios de comunicación con más canales de sindicación			
Prensa			
Rango	Medio	URL	Nº total de canales de sindicación únicos
1	Araraquara.com	http://www.araraquara.com	545
2	UOL o Melhor Conteúdo	http://www.uol.com.br	264
3	Opovo online	http://www.opovo.com.br	230
4	A Tribuna Mato Grosso	http://www.tribunamt.com.br	215
5	Tribuna do Povo	http://www.tribunadopovo.com.br	178
6	Clic RBS	http://www.clicrbs.com.br	129
7	G1 o Portal de noticias da Globo	http://g1.globo.com	96
8	Folha de S.Paulo	http://www.folha.uol.com.br	85
9	Paraná Online	http://www.parana-online.com.br	83
10	Ciencia Hoje	http://www.cienciahoje.uol.com.br	54
Subtotal			1.879
Total recopilado			2.730
Radio			
Rango	Medio	URL	Nº total de canales de sindicación únicos
1	GloboRadio.com	http://globoradio.globo.com	163
2	Stylo 102.1 fm	http://www.stylofm.com.br	160



3	Jovem pan Online	http://jovempan.uol.com.br	110
4	Educativa fm	http://www.educativa.fm.br	55
5	Rádio Gazeta 890 am	http://www.gazetaam.com	50
6	Opovo Rádio	http://www.opovo.com.br	41
7	Universidade ufma 106.9	http://www.universidadefm.ufma.br	38
8	Rádio Campo Aberto	http://www.radiocampoaberto.com.br	27
9	Rádio Cultura 770	http://www.radiocultura770.com.br	19
10	Rádio Guaíba 101.3 fm	http://www.guaiba.com.br	15
Subtotal			678
Total recopilado			1.079
<b>Televisión</b>			
Rango	Medio	URL	Nº total de canales de sindicación únicos
1	Agência de Notícias do Acre	http://www.videos.ac.gov.br	92
2	Jangadeiro TV	http://www.jangadeiroonline.com.br	84
3	Band news TV	http://www.band.com.br	83
4	EBC Brasil TV	http://www.ebc.com.br	61
5	SporTV	http://sportv.globo.com	51
6	TV Novo Tempo	http://novotempo.com	37
7	Canção Nova TV	http://www.cancaonova.com	28
8	IRDEB TV	http://www.irdeb.ba.gov.br	24
9	Rede Globo TV	http://redeglobo.globo.com	23
10	TV Gazeta	http://www.tvgazeta.com.br	9
Subtotal			492
Total recopilado			700

Tabla 8. Medios de comunicación brasileños con más canales de sindicación

Con todo, como se apuntaba en el planteamiento y motivación de la investigación, Google News emplea un sistema de información y agregación centralizado, con filtros avanzados para unificar el acceso a toda la información, con independencia del país y desde una única dirección, en este caso Google News Brasil (GOOGLE, 2012b). La unificación de los medios, evita la dispersión de las consultas y logra que el tráfico de los lectores se centre en una sola herramienta, lo cual plantea una optimización del tiempo que el usuario dedica a la lectura y a la búsqueda de información para su contraste. Ello también explica la negativa de los principales grupos de comunicación brasileños a perder la cuota de mercado, en favor de soluciones técnicas más avanzadas con las que Google logra obtener mejores rendimientos económicos.

También se ha elaborado un análisis de los co-enlaces de los medios de comunicación en prensa, radio y televisión. Siguiendo el método de matriz asimétrica (ORTEGA, J.L.; Aguillo, I., 2006, p.272). La *tabla 9*, muestra el resultado armonizado del co-enlace de todas las páginas de cada sitio web de prensa, radio y televisión. Dada su extensión se han seleccionado los diez medios más co-enlazados de cada sector, obteniendo resultados de interés por consolidar una vez más la presencia de los principales grupos de comunicación y la entrada de otros nuevos como *Diários Associados DA* y su periódico digital *Correio Braziliense*, el grupo *R7* con *radio Guaíba* o el *Grupo Gazeta* que aglutina sectores de prensa, radio y televisión.

<b>Medios de comunicación más coenlazados</b>				
<b>Prensa</b>				
Rango	Medio	URL	Sí coenlazan	No coenlazan
1	G1 Portal da Globo	http://g1.globo.com	2.256	6824
2	Correio Braziliense	http://www.correio braziliense.com.br	921	938
3	Correio Web	http://www.correioweb.com.br	832	317
4	O Sol Diário	http://osoldiario.clicrbs.com.br	761	4832
5	Em. Estado de minas	http://www.em.com.br	669	880
6	UOL – NE10	http://ne10.uol.com.br	619	721
7	UOL – A Tarde	http://www.atarde.com.br	530	542
8	UOL – Ciência Hoje	http://www.cienciahoje.uol.com.br	511	750
9	Gazeta Esportiva	http://www.gazetaesportiva.net	472	325
10	Brazzil Magazine	http://www.brazzil.com/	382	905
<b>Radio</b>				
Rango	Medio	URL	Sí coenlazan	No coenlazan
1	GloboRadio.com	http://globoradio.globo.com	1.647	2.971
2	CBN Cascavel 1240 AM	http://www.cbncascavel.com.br	1.208	296
3	CulturaBrasil	http://www.culturabrasil.com.br	961	3.472
4	Educativa fm	http://www.educativa.fm.br	420	2.259
5	R7 – Radio Guaíba	http://www.guaiba.com.br	405	185
6	UOL – Jovem Pan	http://jovempan.uol.com.br	271	1.335
7	Elo fm	http://www.elofm.com.br	205	426
8	EBC Radios	http://www.ebc.com.br	101	985
9	98 fm	http://www.98fmcuitiba.com.br	100	93

10	Radio Alpha fm	http://www.alphafm.com.br	90	85
Televisión				
Rango	Medio	URL	Sí coenlaza	No coenlaza
1	Band TV	http://band.com.br	4.613	3.267
2	Rede Globo TV	http://redeglobo.globo.com	3.306	4.687
3	SporTV	http://sportv.globo.com	960	3.075
4	Boa Vontade TV	http://www.boavontade.com	430	480
5	Gazeta TV	http://www.tvgazeta.com.br	387	318
6	EBC TV Brasil	http://www.ebc.com.br	345	1.209
7	Boas Novas TV	http://boasnovas.tv	326	154
8	Mix TV	http://www.mixtv.com.br	288	806
9	Globo – Canal Viva	http://canalviva.globo.com	275	1.251
10	Novo Tempo TV	http://novotempo.com/tv/	258	645

Tabla 9. Coenlaces en medios de comunicación de Brasil

Los sitios web más enlazados en los medios de comunicación, pueden ser consultados en la *tabla10*. Se muestra una relación de los diez sitios web más enlazados en la prensa, radio y televisión brasileña. En prensa destacan los sitios del portal de noticias del grupo *Clic RBS*, el diario digital *Zero Hora* y *Araraquara.com*. En radio el sitio de *Calypso fm* del grupo Opovo, *GloboRadio.com* del grupo Globo y *Radio Educativa fm* de la Prefectura Municipal de Campinas. Los sitios web de televisión más enlazados son *Globo TV* y *SporTV* del grupo Globo y *TViG* del grupo de medios iG.

Sitios web más enlazados			
Prensa			
Rango	Medio	URL	Nº de veces enlazado
1	Clic RBS	http://www.clicrbs.com.br	807
2	Zero Hora	http://www.zerohora.com	778
3	Araraquara.com	http://www.araraquara.com	577
4	Diário Gaúcho	http://www.diariogaucha.com.br	523
5	Zero Hora - Donna	http://www.donna.zerohora.com.br	522
6	Donna da Capa	http://www.donnadacapa.com.br	522
7	Fundação Mauricio Sirotsky Sobrinho	http://www.fmss.org.br	522
8	Portal Social	http://www.portalsocial.org.br	522
9	Quem disse	http://www.quemdisse.com.br	522
10	Opovo online	http://www.opovo.com.br	418
Radio			
Rango	Medio	URL	Nº de veces enlazado
1	Calypso fm	http://www.calypsofm.com.br	350
2	GloboRadio.com	http://globoradio.globo.com	333
3	Educativa fm	http://www.educativa.fm.br	204
4	Jovem Pan Radio	http://jovempn.uol.com.br	198
5	EBC Radio	http://www.ebc.com.br	115
6	Radioagência Nacional	http://radioagencianacional.ebc.com.br	112
7	CulturaBrasil	http://www.culturabrasil.com.br	108
8	Nova fm	http://www.novafm103.com.br	75
9	RadarCultura	http://radarcultura.cmais.com.br/	68
10	Radio UNESP	http://radio.unesp.br/	67
Televisión			
Rango	Medio	URL	Nº de veces enlazado
1	Globo TV	http://video.globo.com	601
2	Sport TV	http://sportv.globo.com	330
3	TV IG	http://tvig.ig.com.br	316
4	Univesp TV	http://univesp.tv.br	231
5	Multicultura TV	http://multicultura.cmais.com.br	231
6	Band news TV	http://bandnewstv.band.com.br	175
7	Globosat canais	http://canaisglobosat.globo.com	161
8	Band TV Terra viva	http://tvterraviva.com.br	149
9	Canal viva Globo TV	http://canalviva.globo.com	139
10	Canal Brasil Globo TV	http://canalbrasil.globo.com	135

Tabla 10. Sitios web de medios de comunicación brasileños más enlazados

Finalmente, con toda la información obtenida, el programa Mbot puede ser configurado para generar automáticamente un archivo codificado en lenguaje DOT (GRAPHVIZ, 2012) que contiene todas las relaciones y enlaces obtenidos durante el proceso de webcrawling. Ello permite generar una gráfica aproximada de la estructura y dimensiones de la web de medios de comunicación brasileños. Para poder comparar la prensa, radio y televisión, se han creado gráficas independientes que muestran y corroboran algunas de las afirmaciones realizadas en la investigación. Por ejemplo la *figura1*, correspondiente al mapa web de la prensa digital brasileña, permite observar una estructura altamente ramificada en los extremos de la elipse.

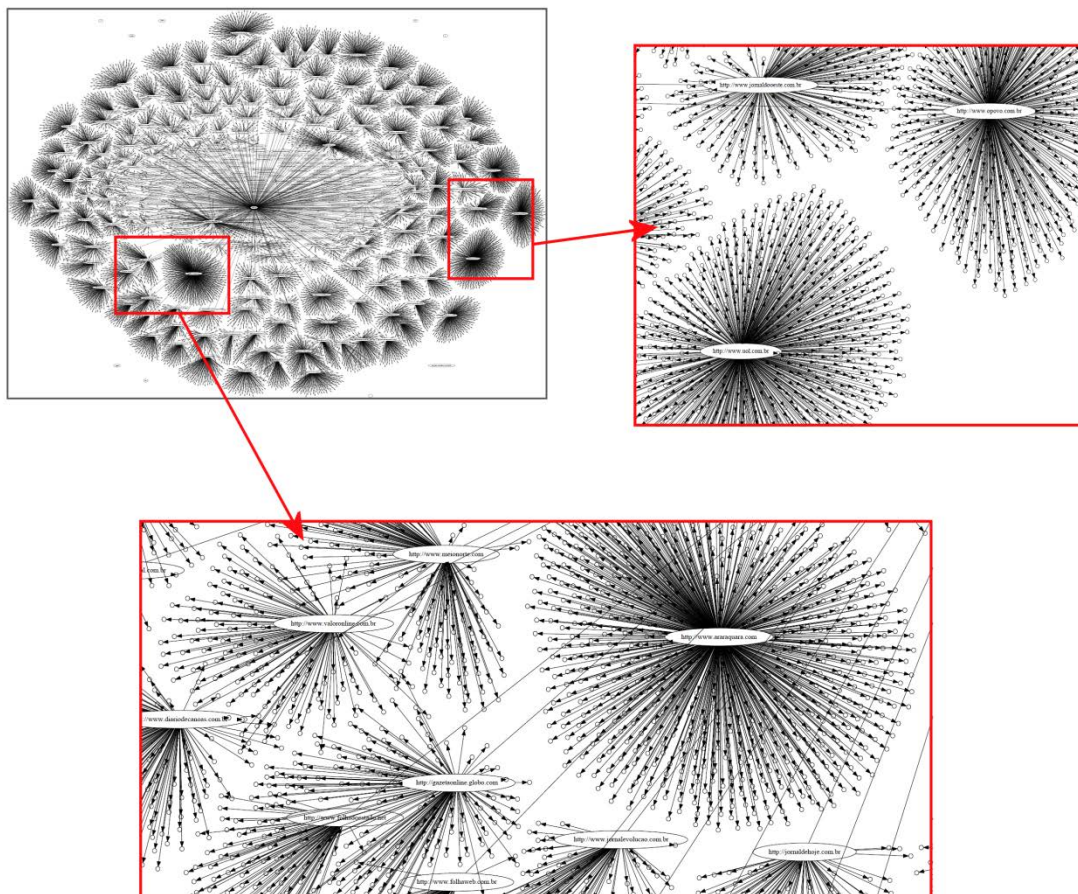


Figura 1. Representación gráfica de la web de prensa brasileña

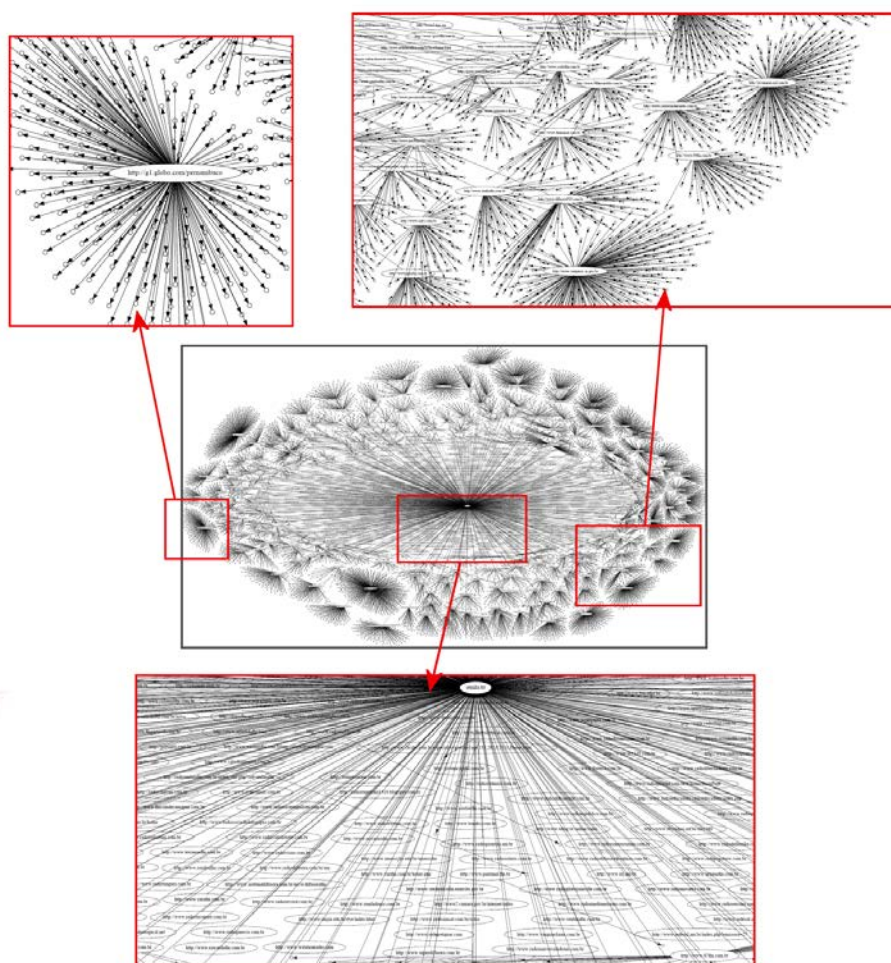
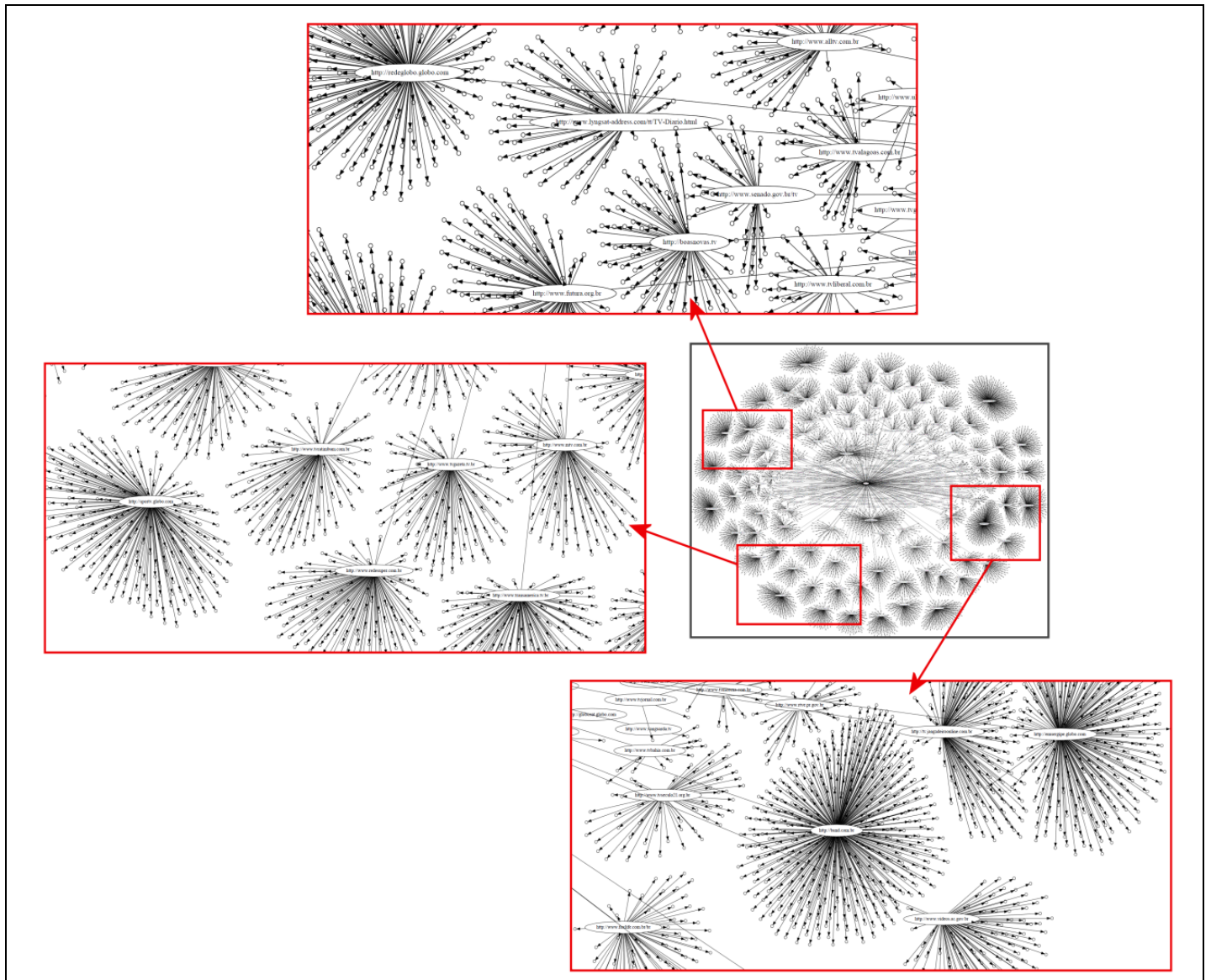


Figura 2. Representación gráfica de la web de radio brasileña



Los extremos más ramificados corresponden a los principales grupos de comunicación, tal como puede advertirse en las ampliaciones de la imagen. También se observa que existe un núcleo u origen en el que confluyen todas las relaciones. Esto significa, que es la semilla original, a partir de la que se origina todo el estudio. De esta forma, es fácil apreciar cuál es el sitio web más desarrollado y mejor enlazado con respecto a todos los demás. Otro detalle significativo es la disposición de una serie de enlaces no ramificados en el centro de la gráfica, formando una elipse dentro de otra elipse. Estos enlaces corresponden a la componente de la web menos vinculada, también denominada *Islands*.

En la *figura2*, se muestra el mapa de la web correspondiente a la radio brasileña. Su estructura está menos ramificada que la prensa y hacia el centro de la elipse se observa un mayor número de enlaces pobremente vinculados. Esto se debe al menor número de enlaces únicos por sitio web en la radio en comparación con la prensa, unido a una gran cantidad de medios de radio, 734 en relación a los de prensa 340.



*Figura 3. Representación gráfica de la web de televisión brasileña*

En la *figura3*, se muestra el mapa web de la televisión brasileña, con una estructura más proporcionada, parecida al mapa web de la prensa. No obstante, al igual que en el caso de la radio, consta de una amplia elipse interna que contiene un 89% de páginas débilmente vinculadas con el resto.

#### 4. Conclusiones y perspectiva

1. Los principales grupos de comunicación brasileños, dada su reiterada presencia en el análisis webmétrico realizado son el grupo UOL, Globo, Clic RBS, Opovo, iG, Gazeta, DA y R7.
2. Los medios de comunicación brasileños constan de 2 millones de enlaces de los cuáles son únicos sólo medio millón. La prensa ocupa un 42% del total (204.133 enlaces), la radio el 28% (133.322 enlaces) y la televisión el 30% (146.257 enlaces). Ello indica que el sector predominante es la prensa con un número menor de medios que la radio (54% menos) y mayor que la televisión (29% más).
3. La web de medios de comunicación brasileños tiene una componente de páginas web estáticas del 71% frente a páginas web dinámicas del 29%, lo que significa que gran parte de las noticias que publican los medios de prensa se elabora en *.html* frente al empleo de formatos de tipo *.pdf* o *.doc*, que resultan minoritarios.
4. La componente de enlaces fuertemente vinculados en la web de prensa, radio y televisión brasileña representa entre el 7,7% y el 10,95% de todos los enlaces analizados, indicando una pobre interconexión entre el 80% de los enlaces restantes. Ello es debido en parte a la desigualdad entre los grandes grupos de comunicación que representan la sección más enlazada y los medios de menor tamaño. También se ha demostrado que la gran cantidad de información y noticias es publicada en páginas web estáticas lo que sugiere un nivel de enlazamiento menor que los sitios web principales o portales de información.
5. El empleo de metadatos está muy extendido en los principales grupos de comunicación brasileños. Los primeros diez medios de prensa con más metadatos aglutinan el 41% de todos los metadatos del sector de prensa, lo que indica que pueden existir grandes diferencias de meta-descripción entre los distintos medios de comunicación, así como en los gestores de contenidos que se emplean.
6. Los medios de comunicación con mayor tasa de coenlaces son el portal de noticias del grupo Globo (2.256 enlaces), GloboRadio.com (1.647 enlaces) y Band TV del grupo UOL (4.613 enlaces). En el sector de televisión, se aprecia además que Globo TV es la segunda web de televisión más coenlazada, lo que significa que el grupo Globo es el más coenlazado de Brasil.
7. El análisis de los medios de comunicación brasileños en la web ha dado como resultado la extracción de 4.509 canales de sindicación, unos 3.000 canales más que los empleados en el portal de noticias Google News Brasil. Esto significa que el seguimiento de la producción informativa y de contenidos en la web de prensa, radio y televisión puede ser mejor estudiada, dada la exhaustividad en la recuperación de las fuentes de información. Ello posibilitará el próximo estudio sobre la producción informativa de los medios de comunicación brasileños y su categorización temática.
8. Teniendo en cuenta la ruptura de los principales medios de comunicación brasileños con el agregador Google News, se propone el desarrollo de un portal de noticias brasileño que contemple todos los canales de sindicación de los medios de comunicación recopilados, de forma tal que actúe como portal integrador de noticias de prensa, radio y televisión, con un modelo de redirección del tráfico a las fuentes de información original. Éste modelo podría ser desarrollado en cooperación con los principales organismos e instituciones de comunicación del país, representando un ejemplo de democratización de la información y un instrumento de ecuanimidad entre grandes y pequeños grupos de comunicación.

## 5. Referencias

- ANDREWS, R. 2012. *Google News faces mass newspaper boycott in Brazil*. En: *PaidContent*. [online]. [Consulta 5 Nov 2012]. Disponible en: <http://paidcontent.org/2012/10/19/google-news-faces-mass-newspaper-boycott-in-brazil/>
- ARANTES FROTA, M.H.; D.N. BARBOSA DE SOUZA. 2005. Mass Media and Empowerment in Brazil: The Loss of Innocence. In: Anne Marie BÜLOW-MØLLER, (ed). *Proceedings of The Association for Business Communication 7th European Convention*. Copenhagen: Association for business communication, pp.1-10.
- BAEZA YATES, R., LÓPEZ, V.; C. CASTILLO. 2006. Características de la Web de España. *El Profesional de la Información*. **15**(1), pp.6-17.
- BBC. 2012. *Brazilian newspapers pull out of Google News*. [online]. [Consulta 05 Nov 2012]. Disponible en: <http://www.bbc.co.uk/news/world-latin-america-20018221>
- BLÁZQUEZ OCHANDO, M. 2011. *Primeras pruebas del mbot webcrawler*. [online]. [Consulta 2 Oct 2012]. Disponible en: <http://www.mblazquez.es/documents/articulo-pruebas1-mbot.html>
- BLÁZQUEZ OCHANDO, M. 2012. Desarrollo de un sistema de clasificación automática de contenidos en medios de comunicación hispano-mexicanos. In: *9º Seminario Hispano Mexicano de Bibliotecología y Documentación (7-9 mayo)*. México DF: UNAM.
- BLÁZQUEZ OCHANDO, M.; SERRANO MASCARAQUE, E. 2011. Análisis de la web y usabilidad: prueba de funcionamiento de Mbot webcrawler. In: *X Congreso ISKO Capítulo Español (Ferrol, 30 junio - 1 julio 2011)*. Ferrol: ISKO.
- GOOGLE. 2012a. *Google Notícias Brasil: Sobre o Google Notícias*. [online]. [Consulta 10 Nov 2012]. Disponible en: [http://news.google.com.br/intl/pt-BR\\_br/about\\_google\\_news.html](http://news.google.com.br/intl/pt-BR_br/about_google_news.html)
- GOOGLE. 2012b. *Google Notícias Brasil*. [online]. [Consulta 10 Nov 2012]. Disponible en: <http://news.google.com.br>
- GRAELLS, E.; BAEZA YATES, R. 2007. *Características de la Web Chilena 2007*. Santiago de Chile.
- GRAPHVIZ. 2012. *The DOT Language*. [online]. [Consulta 1 Oct 2012]. Disponible en: <http://www.graphviz.org/content/dot-language>
- HONKENSON, T.C. 2012. *Radio Station World: List of Brazil stations on the Web*. [online]. [Consulta 10 Nov 2012]. Disponible en: [http://radiostationworld.com/locations/brazil/radio\\_websites.asp](http://radiostationworld.com/locations/brazil/radio_websites.asp)
- INDVIK, L. 2012. *Major Brazilian Newspapers Withdraw From Google News*. En: *Mashable Business*. [online]. [Consulta 5 Nov 2012]. Disponible en: <http://mashable.com/2012/10/19/brazilian-newspapers-google-news/>
- IVC. 2012. *Instituto Verificador de Circulação*. [online]. [Consulta 10 Nov 1012]. Disponible en: <http://www.ivcbrasil.org.br/>
- MEDIA ONLINE. 2012. *Media Brazil*. [online]. [Consulta 1 Nov 2012]. Disponible en: <http://www.mediaonline.net/en/brazil>
- ORTEGA, J.L.; AGUILLO, I. 2006. Análisis de co-enlaces: una aproximación teórica. *El Profesional de la Información*. **15**(4), pp.270-277.

RALPH, T. 2012. *Brazil's newspapers opt out of Google News*. En: *Globalpost*. [online]. [Consulta 5 Nov 2012]. Disponible en: <http://www.globalpost.com/dispatch/news/regions/americas/brazil/121021/brazils-newspapers-opt-out-google-news>

WEBSTER, A. 2012. *Brazilian newspapers yank content from Google News, blame lack of compensation*. En: *The Verge*. [online]. [Consulta 5 Nov 2012]. Disponible en: <http://www.theverge.com/2012/10/21/3534782/brazilian-newspapers-boycott-google-news>